"Machine Learning and Computational Statistics"

5th Homework

Exercise 1:

Consider the Erlang distribution $p(x) = \theta^2 x \exp(-\theta x)u(x)$, (where u(x)=1(0), if $x \ge 0$ (<0)), whose mean equals to $2/\theta$.

- (a) Given a set of N measurements $x_1,...,x_N$, for the random variable x that follows the Erlang distribution, prove that the ML estimate of θ is $\vartheta_{\text{ML}} = 2N / \sum_{i=1}^{N} x_i$
- (b) For N=5 and $x_1=2$, $x_2=2.2$, $x_3=2.7$, $x_4=2.4$, $x_5=2.6$, estimate the mean of the random variable x.

Exercise 2:

Consider again the Erlang distribution $p(x)=\theta^2 x \exp(-\theta x)u(x)$, (where u(x)=1(0), if $x\geq 0$ (<0)). Given

- a set of N measurements $x_1,...,x_N$, for the random variable x that follows the Erlang distribution, and
- the a priori probability for the parameter θ is a normal distribution, $N(\theta_0, \sigma_0^2)$ (where θ_0, σ_0^2 are known)
- (a) Compute the MAP estimate of the parameter θ .
- **(b)** How this estimate becomes for the case were (i) N $\rightarrow \infty$, (ii) $\sigma_0^2 >>$ and (c) $\sigma_0^2 <<$? Give a short justification.

Exercise 3 (python code + text):

Suppose we have a copper wire and we measure (experimentally) its resistance at various temperatures. An Nx2 array, called Data, of relative measurements is given in the file HW5.mat. Each row of the array corresponds to a measurement (θ_i, R_i) , i = 1, ..., N, where θ_i corresponds to the temperature and R_i to the associated resistance.

- (i) Determine the relation $R = f(\theta)$, assuming that the joint pdf of the random variables θ and R, corresponding to the temperature and the resistance, respectively, is normal and utilizing the MSE criterion.
- (ii) Use the Data_test matrix to assess the performance of the regressor. Specifically, considering each row of the matrix, use the first element (temperature) and

estimate the value of the resistance using the relation derived in (i). Then, compare this estimate with second element of the row (real value).